LEVANTAMENTO E MAPEAMENTO DA DISTRIBUIÇÃO DE MIDDLEBOXES PRESENTES NA WORLD WIDE WEB

Bruno Chagas[1]
Adenes Schwantz[2]

INTRODUÇÃO

Middleboxes são dispositivos presentes em redes de computadores que transformam, inspecionam, filtram ou, de qualquer outra forma, manipulam e modificam o tráfego de pacotes para diferentes propósitos do que roteamento (CARPENTER e BRIM).

Estes dispositivos estão fortemente presentes em redes empresariais com propósito de aprimoramento da segurança e desempenho da rede. Esta larga aplicação de *middleboxes* resulta em vários desafios e muito criticismo relativos à pobre interação das mesmas com as camadas superiores dos protocolos de rede (SHERRY, HASAN e SCOTT).

Essa pesquisa tem por objetivo, através de uma análise concentrada, em diversos caminhos de rede, levantar os tipos mais comuns de *middleboxes* presentes, envolvendo a maior amostra possível de endereços IP da *word wide web*, identificar o tipo de *middlebox* que causa as modificações encontradas e sua eventual necessidade, para assim levantar sua distribuição, justificativa de uso, e possíveis efeitos nocivos. Além de validar ferramentas novas e recentemente propostas, essa investigação ainda visa ampliar e contribuir de forma significativa para aumentar a base de informação e conhecimento no que se refere ao desempenho e interferência dos mais diversos aparatos instalados (quase que descontroladamente) nas redes de computadores atuais.

MATERIAL E MÉTODOS

Este projeto fará uso da ferramenta “Tracebox” para mapear *middleboxes* e analisar as modificações ocorridas nos pacotes em cada uma delas.

Primeiramente, foi levantada uma lista de endereços, tais que estão distribuídos em todos os continentes. Para cada endereço é enviado três pacotes (utilizando três métodos diferentes, pacotes estes chamados de *probes*) para a detecção das *middleboxes*.

Os testes geram uma quantidade significativa de dados, pois foram realizadas em São Paulo/SP e em Videira/SC utilizando conexão Ethernet e Wifi, somando aproximadamente quatro mil pacotes enviados.

[1]Aluno do Instituto Federal Catarinense, Videira, SC. Curso Superior de Engenharia Elétrica. E-mail: abschagas@hotmail.com
[2]Professor Orientador do Instituto Federal Catarinense, Videira, SC. Curso Superior de Engenharia Elétrica. E-mail: adenes.schwantz@ifc-videira.edu.br

Para a compilação dos dados é utilizado um código desenvolvido em linguagem de programação python que contabiliza as mudanças ocorridas no destino final e posteriormente mostra o número de vezes que cada mudança ocorreu. Futuramente este código fará uma análise ainda mais profunda, contabilizando mudanças em todas os *hops*.

Os dispositivos mais presentes nas detecções serão estudados, a partir de qual ponto sua necessidade é tão essencial e se poderiam ser substituídos ou não. Os tempos de comunicação entre fonte e destino serão de fundamental importância, e serão comparados caminhos livres com caminhos com interferência.

Ao final, todos os dados serão disponibilizados publicamente, sendo os mais relevantes discutidos, analisados e poderão vir acompanhados de modificações sugeridas.

RESULTADOS E DISCUSSÃO

Em sua fase inicial, o projeto ainda não conta com uma análise profunda dos resultados coletados. Porém, já conta com um mecanismo para facilitar a contagem das modificações encontradas, que, como pode ser observado na Figura 1 seria pouco prático realizar este trabalho manualmente.

Figura 1- Resultado do envio de uma *probe*.

```
traceboxto 103.252.142.21 (www.irctc.co.in): 64 hops max
1: 192.168.0.1 0ms  [PARTIAL]
2: 10.32.92.1 8ms  [PARTIAL] IP::TTL IP::CheckSum
3: 201.6.33.193 9ms  [PARTIAL] IP::TTL IP::CheckSum
4: 201.6.40.197 8ms  [PARTIAL] IP::TTL IP::CheckSum
5: 201.6.40.209 9ms  [PARTIAL] IP::TTL IP::CheckSum
6: 201.6.42.14 10ms  [PARTIAL] IP::TTL IP::CheckSum
7: 201.6.42.13 12ms  [PARTIAL] IP::TTL IP::CheckSum
8: 201.90.40.25 11ms  [PARTIAL] IP::TTL IP::CheckSum
9: 200.244.213.132 14ms  [PARTIAL] TCP::SeqNumber IP::TTL IP::CheckSum
10: 200.230.220.174 123ms  [PARTIAL] IP::TTL IP::CheckSum
11: 129.250.202.185 122ms  [PARTIAL] IP::TTL IP::CheckSum
12: 129.250.5.61 125ms IP::TTL IP::CheckSum +RawLayer
13: 129.250.4.13 198ms IP::TTL IP::CheckSum +RawLayer
14: 129.250.2.54 193ms IP::TTL IP::CheckSum +RawLayer
15: 129.250.3.124 9208ms IP::TTL IP::CheckSum +RawLayer
16: 129.250.2.102 207ms IP::TTL IP::CheckSum +RawLayer
17: 129.250.3.49 390ms IP::TTL IP::CheckSum +RawLayer
18: 129.250.3.147 393ms  [PARTIAL] IP::TTL IP::CheckSum
19: 129.250.12.230 385ms  [PARTIAL] IP::TTL IP::CheckSum
20: *
21: 125.19.0.142 406ms  [PARTIAL] IP::TTL IP::CheckSum
22: 103.252.142.200 415ms TCP::CheckSum IP::TotalLength IP::TTL IP::CheckSum -TCPOptionMPTCPCapable +RawLayer
23: 103.252.142.21 405ms TCP::SrcPort TCP::DstPort TCP::SeqNumber TCP::AckNumber TCP::DataOffset TCP::Flags TCP::WindowsSize TCP::CheckSum
IP::TotalLength IP::Identification IP::TTL IP::CheckSum IP::SourceIP IP::DestinationIP -TCPOptionTimestamp -TCPOptionPad -
TCPOptionSACKPermitted -TCPOptionMPTCPCapable -TCPOptionWindowScale
```

Fonte: Construção dos autores.

Cada lista pode conter centenas de arquivos como a Figura 1, porém, com a utilização do mecanismo, pode-se chegar a um resultado como mostrado na Figura 2, feito no formato xlsx (pode ser executado com o *software* Microsoft Excel ou libreoffice).

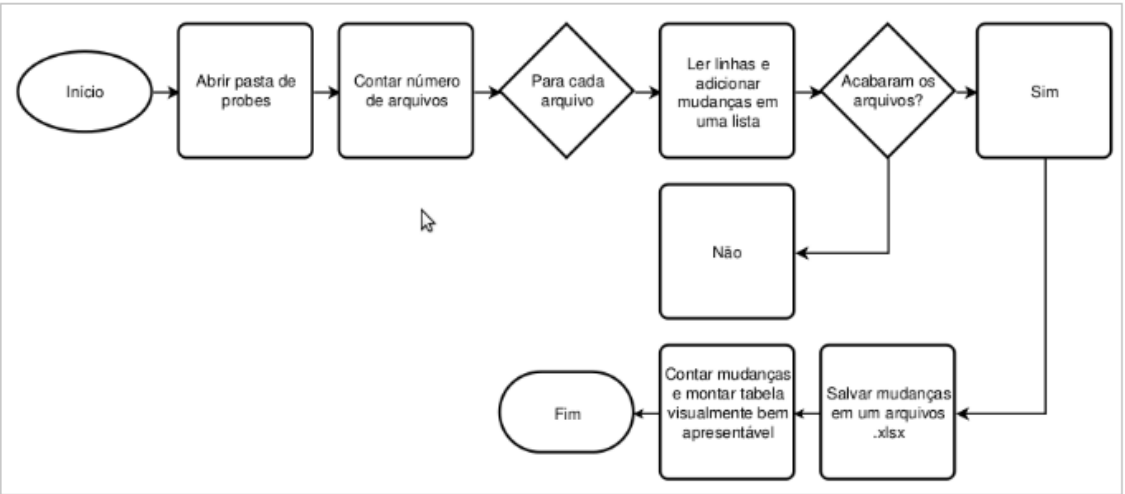
Figura 2 - Contagem dos resultados encontrados em uma das listas.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	TCP	Count	IP	Count	TCPOptio	Count	TCPOptio	Count	Removido	Count	TCPOptio	Count	Adicionad	Count	
2	SrcPort	73	TotalLeng	73	Value	54	Shift	63	-TCPOptio	73	MaxSegS	11	+RawLay	3	
3	DstPort	73	Identifica	73	EchoRepl	53			-TCPOptio	19					
4	SeqNumb	73	TTL	78					-TCPOptio	4					
5	AckNumb	73	CheckSur	78					-TCPOptio	6					
6	DataOffse	73	SourceIP	73					-TCPOptio	10					
7	Flags	73	Destinati	73					-TCPOptio	2					
8	Windows	73	Flags	2											
9	CheckSur	75	DiffServio	1											
10															

Fonte: construção dos autores.

Ainda em desenvolvimento, somente as modificações ocorridas no destino final são contabilizadas. Porém, será criado um método para facilitar a visualização e análise das modificações encontradas em cada *middlebox*, pois em formato de texto, a análise levará muito tempo, como mostrado na Figura 3.

Figura 3 - Fluxograma para a analisar resultados de uma lista.



Fonte: construção dos autores.

Futuramente, será desenvolvida uma técnica para somar os resultados automaticamente, semelhante ao mostrado na Figura 2, eliminando a necessidade de uma tarefa manual e facilitando assim o andamento do projeto.

Contabilizados os dados das maiores listas, somando 296 *probes*, realizadas com conexão Ethernet a partir de São Paulo/SP encontra-se os resultados mostrados na Tabela 1. A ocorrência de alguns dos itens, como por exemplo IP-TTL e IP-CheckSum, é normal, devido a uma característica do próprio pacote, pois estes são itens verificadores e sua mudança ocorreria mesmo sem a presença de uma middlebox.

Tabela 1 – Maiores ocorrências nas listas

TCP		%	IP		%	Removidos		%
SrcPort	248	83.78	TotalLength	244	82,43	TCPOptionMPTCPCapable	244	82,43
DstPort	248	83,78	Identification	239	80,74	TCPOptionTimestamp	76	25,68
SeqNumber	248	83,78	TTL	262	88,51	TCPOptionPad	11	3,72
AckNumber	248	83,78	CheckSum	262	88,51	TCPOptionSACKPermitted	19	6,42
DataOffset	239	80,74	SourceIP	248	83,78	TCPOptionWindowScale	36	12,16
Flags	248	83,78	DestinationIP	248	83,78	TCPOptionMaxSegSize	4	1,35
WindowsSize	244	82,43	DiffServicesCP	3	1,01			
CheckSum	251	84,80	Flags	16	5,41			

Fonte: Construção dos autores

É possível perceber que há um número muito grande de ocorrências na maioria dos itens. Contudo, há itens, como *DiffServicesCP* e *TCPOptionMaxSegSize*, que tiveram respectivamente 1,01% e 1,35% dos resultados.

O local de onde essas mudanças ocorreram podem ser facilmente encontradas e futuramente estudadas com mais facilidade.

CONSIDERAÇÕES FINAIS

Hodiernamente, a maioria das pessoas utiliza a internet diariamente, executando todo o exposto aqui e desenvolvido na atividade de pesquisa espera-se alcançar resultados sólidos que evidenciem a geolocalização de *middleboxes*, sua distribuição, seus efeitos e potencial influência no tráfego da grande rede.

REFERÊNCIAS

CARPENTER B.; BRIM S., **Middleboxes: Taxonomy and issues**. Internet Engineering Task Force, RFC 3234, 2002.

SHERRY J.; HASAN S.; SCOTT C.; KRISHNAMURTHY A.; RATNASAMY S.; SEKAR V. **Making Middleboxes Someone Else's Problem: Network Processing as a Cloud Service**. In: Proc. ACM SIGCOMM Conference, v. 18 n. 85 p. 13-24, 2012.